Cytometry Part A 71A:439–450 (2007)

Review Article

Automated Analysis of FISH and Immunohistochemistry Images: A Review

Zenonas Theodosiou,[1] Ioannis N. Kasampalidis,[1] George Livanos,[2] Michalis Zervakis,[2] Ioannis Pitas,[1]* and Kleoniki Lyroudia[3]

[1]Department of Informatics, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece
[2]Department of Electronic and Computer Engineering, Technical University of Crete, University Campus, Kounoupidiana, 73100 Chania, Greece
[3]Department of Endodontology, School of Dentistry, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece

Received 14 December 2006; Accepted 13 March 2007

Fluorescent in-situ hybridization (FISH) and immunohistochemistry (IHC) constitute a pair of complimentary techniques for detecting gene amplification and overexpression, respectively. The advantages of IHC include relatively cheap materials and high sample durability, while FISH is the more accurate and reproducible method. Evaluation of FISH and IHC images is still largely performed manually, with automated or semiautomated techniques increasing in popularity. Here, we provide a comprehensive review of a number of (semi-) automated FISH and IHC image processing systems, focusing on the algorithmic aspects of each technique. Our review verifies the increasingly important role of such methods in FISH and IHC; however, manual intervention is still necessary in order to resolve particularly challenging or ambiguous cases. In addition, large-scale validation is required in order for these systems to enter standard clinical practice. © 2007 International Society for Analytical Cytology

Key terms: fluorescent in situ hybridization (FISH); immunohistochemistry (IHC); image analysis techniques; nuclei segmentation; spot detection; antibody staining

A variety of methods are available for the detection of gene status in tissue samples, with fluorescence in situ hybridization (FISH) and immunohistochemistry (IHC) being two of the most prominent ones for detecting gene amplification and overexpression, respectively. Both techniques permit the study of small amounts of formalin-fixed, paraffin-embedded tissue and the interpretation of the findings on a cell-by-cell basis. FISH allows selective staining of various DNA sequences with fluorescent markers, and thereby the detection, analysis, and quantification of specific numerical and structural abnormalities within nuclei. This procedure has proven to be as accurate as Southern blot analysis, while allowing the measurement of the fraction of amplified cells and the intercellular heterogeneity within a given cell population (1,2). On the other hand, IHC uses specific antibodies to stain proteins in situ, which allows the identification of many cell types that could be visualized by classical microscopy.

FISH is a direct in situ technique that is relatively rapid and sensitive. No cell culture is needed to apply this method and results are easier to interpret than karyotype. The FISH technique has the advantages of a more objective scoring system and the presence of a built-in internal control consisting of the two Her-2/neu gene signals present in all nonneoplastic cells of the specimen. Disadvantages of FISH testing include the high cost of each test, long time needed for slide scoring, requirement for a fluorescence microscope, inability to preserve the acquired sample for storage and review, and, occasional difficulty in identifying the invasive tumor cells (3). On the other hand, advantages of IHC testing include its wide availability, relatively low cost, easy and long preservation of stained slides, and use of an ordinary light microscope. Disadvantages of IHC include the impact of preanalytic issues including storage, duration and nature of system control samples, and most importantly, the difficulties in applying a subjective slide scoring system (3).

Grant sponsor: eHealthcare, Network of Excellence; Grant number: 508803

*Correspondence to: Professor Ioannis Pitas, Department of Informatics, Box 451, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece.
E-mail: pitas@aiia.csd.auth.gr

Published online 8 June 2007 in Wiley InterScience (www.interscience.wiley.com).
DOI: 10.1002/cyto.a.20409

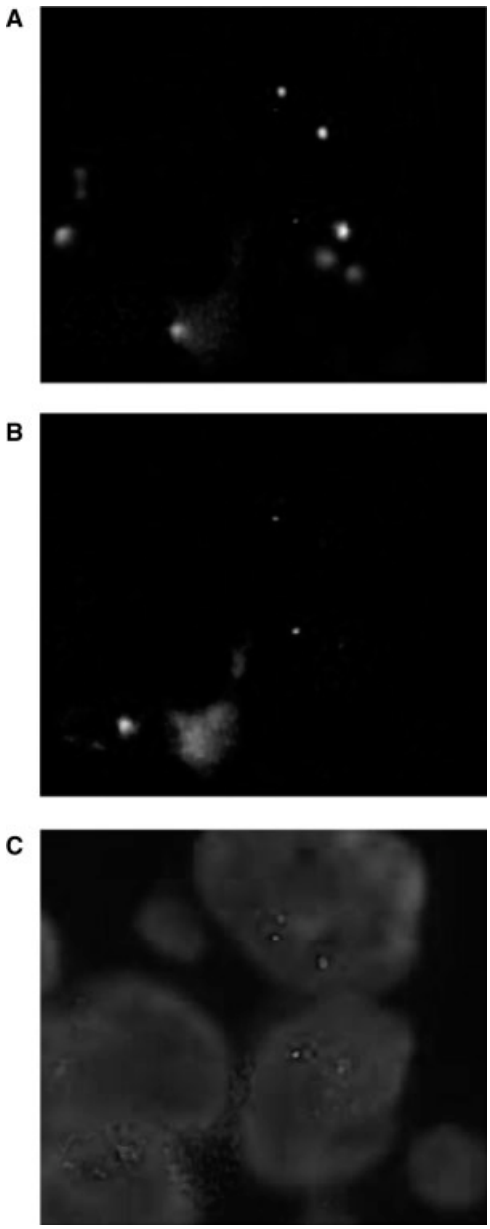

FIG. 1. Example of FISH image from breast tissue sample with FISH probes targeting the Her-2/neu gene and the chromosome 17 centromere. (A) The copies of Her-2/neu gene are represented in red channel. (B) The centromeres of 17 chromosomes are represented in green channel. (C) Nuclei are represented in blue channel. Reprinted from Theodosiou et al., Fish Image Analysis System for Breast Cancer Studies, European Conference on Emergent Aspects in Clinical Data Analysis (EACDA, 2005), Pisa, Italy, September 28-30, 2005.

Evaluation of gene status from FISH images is based on the manual counting of gene signals in interphase nuclei, which become visible as colored dots. As an example, the FDA approved PathVision Her2 FISH kit (Vysis, Downers Grove, USA) uses DNA probes that when applied to a tumor tissue sample target the Her-2/neu gene and attach themselves to its target sequence in a process called hybridization. The probes carry special fluorescent markers that emit a reddish light under a fluorescent microscope, as shown in Figure 1. Similarly, probes for centromere 17 (CEP-17), the chromosome on which the gene Her-2/neu is located, are visible as green spots. Additionally, the sections are counterstained with 4′-6-Diamidino-2-phenylindole (DAPI), providing a blue background for the nucleus body. According to PathVysion Her-2/Neu DNA Probe Kit (PathVysion Kit), evaluation of FISH images involves enumeration of 20 interphase nuclei from tumor cells per target are reported as the ratio of average Her-2/neu copy number to that of CEP-17, where a ratio of Her-2/neu to CEP-17 copy number greater than 2 denotes amplification, while results at or near the cut-off point (1.8–2.2) should be interpreted with caution.

On the other hand, IHC testing reveals the amount of protein product in the cell by using antibody staining, which is usually manifested in brown color. In most cases, samples have also been prestained with hematoxylin-eosine which facilitates tissue examination. The effect of this staining is the coloring of the nucleus as blue and the rest of the tissue as pink, as shown in Figure 2. The percentage of tumor cells that have completely stained membranes and the intensity of this staining are used by a pathologist to score and classify the result as positive or negative (3). As an example, Ross et al. (3) have established a standardized immunohistochemical procedure and scoring system for evaluating the status of the breast tissue samples stained for Her-2/neu, in which cells containing less than 20,000 receptors would show no staining and are given a score of 0, cells containing less than ~100,000 receptors would show partial membrane staining with less than 10% of the cells showing complete membrane staining and are given a score of 1+, cells containing ~500,000 receptors would show light to moderate complete membrane staining in more than 10% of the cells and are given a score of 2+, while cells containing ~2,300,000 receptors would show strong, complete membrane staining in more than 10% of the cells and are given a score of 3+.

In practice, current analysis of both FISH and IHC images is performed in a semiautomated way, with the aid of image processing software. A study by Klijanienko et al. (4) has shown strong correlation of detection results using

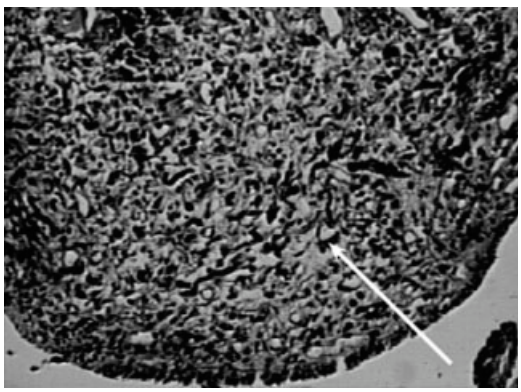

FIG. 2. Micrograph of an inflamed dental pulp tissue immunohistochemically stained with S100B; the white arrow shows a positive neural element. Figure provided by Dr. Dourou, Aristotle University of Thessaloniki.

visual-only and semiautomated methods for evaluating the status of Her-2/neu in breast carcinomas samples.

ANALYSIS OF FISH IMAGES

We have examined a plurality of FISH image analysis methods, the most notable of which are presented here. Fernandez et al. (5) employed the histogram of images which showed two nonoverlapping peaks, one corresponding to the counterstain intensity range 100 and 150 and another one corresponding to the FISH signal intensity range between 175 and 255. Thus, using a simple threshold, segmentation of the FISH signals, referred to as spots or dots, was obtained. Two different parameters, surface, defined as the area in pixels of the binary image and grey mean, defined as the mean of the grey levels present in the segmented region, were used to quantify the amount of positive signal of the hybridization. This serves as an indirect measure of the number of copies of the Her-2/neu gene. Their results show that the number of copies discriminated with the naked eye in Type 2 cells is extremely close to that obtained by the automated method.

Netten et al. (6,7) developed an automated method for counting dots per cell nucleus in slides of lymphocytes from cultured blood. The system contains all the components common to image processing and image analysis: automated focusing, image acquisition, segmentation, measurement, and classification. Image acquisition is followed by an image processing algorithm that actually counts the number of dots. The algorithm is divided into four steps: (1) find a region that contains a nucleus, (2) find the nucleus in the region, (3) find spots in the nucleus, and (4) count spots and update the spot histogram for the entire microscope slide.

The region of interest is identified as follows: the original image is sub-sampled by a factor of eight and the reduced image is prefiltered to suppress noise. Gray-level opening is performed to remove the dots and correct the shades. The resulting image is segmented by an automatically chosen constant threshold, which is suitable for images that contain only a few objects and have a large background area. The region of interest is defined by an enclosing rectangle for each object.

For each ROI, the authors processed the original image at full resolution to define a mask for the nucleus. Then, gray-value opening is applied to remove the dots and the ISODATA threshold algorithm (8) is used to segment the ROI into object and background. The resulting object mask is then further processed using binary morphological operations to remove small objects and to separate slightly overlapping nuclei (9). After segmentation, size, shape, and intensity features were measured for each object (10). The features were used to select single nuclei and to reject touching nuclei, debris, etc.

To detect the FISH signals, the authors segmented again the original images within the mask of the nucleus by employing three different techniques based on the top-hat transform. At the beginning, the top-hat transform is applied on a 5×5 window within the mask of the nucleus on the original image, removing the DAPI counter stain. The dots are found by applying a constant threshold on the top-hat transform, where the constant is defined by

$$\theta_{th} = \mu_{bkg} + k \times \sigma_{bkg} \quad (1)$$

where μ_{bkg},σ_{bkg} are the mean and standard deviation of the background inside the mask of the nucleus and the mean and standard deviation were estimated using the pixels below 90% intensity of the top-hat image. Determination of parameter k was based on a limited number of nuclei used as a training set. After applying the top-hat threshold, most spots were detected but some appeared merged. Therefore, a nonlinear Laplacian filter was applied within the mask of the top-hat threshold and, by using a threshold determined by half the minimum intensity of the Laplacian image, the touching dots were split.

Since the definition of a proper threshold level for the top-hat threshold is a difficult task, the authors also employed a variable threshold level. Pixels with intensity equal to a threshold level are assigned to a dot, if they are connected to this dot. In case they are not connected to an existing dot, a new dot is created. The threshold level θ_{seed} starts at the maximum intensity of the image I_{max} and runs down until it is just above the background level $\theta_{seed} = \mu_{bkg} + k \times \sigma_{bkg}$.

To test the performance of the system, the authors compared the results with manual counting. The number of dots detected by the system is usually larger than the number of dots detected manually and has a higher variance. Moreover, some errors occurred during the analysis like: false dots, missed dots, split dots, overlapping dots, out of focus, and debris. Results suggest that the slide quality has an influence on the system performance, while debris, high autofluorescence, and low probe intensity can make results unreliable.

A system for automatic detection and scoring of FISH signals in interphase nuclei was developed by Solorzano et al. (11). The system performs sequentially the following set of actions: (1) performing all the stage movements and filter stages required to scan the area under study, (2) focusing the microscope on every field of view, (3) acquiring the counter-stained and FISH labeled images, and (4) analyzing the images in order to define the position of the nuclei and the number of FISH signals inside of them.

Once the images are focused, three images, the DNA counterstained image and both FISH signal images, are acquired. Before the images are analyzed, a preprocessing step involving shading correction, background subtraction, autofluorescence correction, and color-shift compensation (12) are used. After correction, images are segmented to extract the required information. Nuclei are segmented by automatically thresholding the histograms of the DNA counterstained images using the ISODATA algorithm. After the distance transform is applied, clusters of nuclei are divided into their individual components using the morphological watershed algorithm (13). FISH signals are extracted by means of the top-hat transform (14) followed by a recursive reconstruction algorithm (15) which

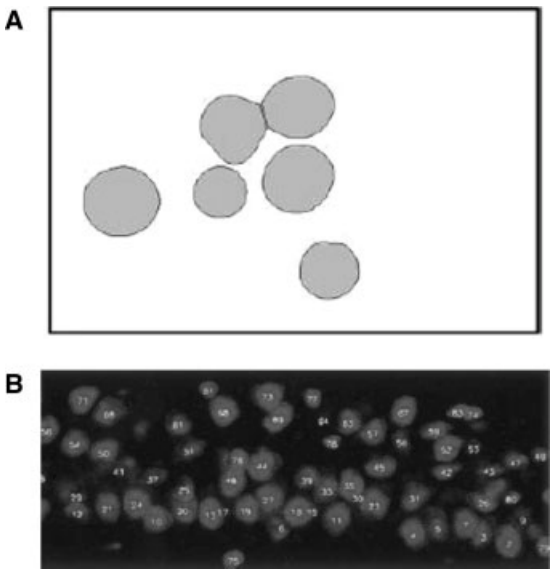

FIG. 3. Nuclei segmentation. (A) Image taken from Kozubek et al. (16). Detection of nuclei within the counterstain image, the final result obtained after the segmentation procedure. Reprinted from Kozubek et al. (16), with permission from Wiley-Liss, Inc., a subsidiary of John Wiley & Sons, Inc. (B). The segmentation after using the algorithm with enhanced 3D watershed segmentation followed by model-based merging. The labeled detected nuclei are presented with red color. Reprinted from Chawla et al. (24), with permission from Elsevier.

removes secondary peaks and refines the contours of the real FISH signals. By evaluating the system, the authors mentioned that the results are comparable to the ones obtained by human experts. Moreover, the use of the internal, nucleus by nucleus, control, provided by the simultaneous use of two probes, improves system sensitivity.

Kozubek et al. (16) developed a completely automated system that acquires and analyses 2D and 3D FISH images. The 3D images were used to study the spatial chromatin arrangement in cell nuclei. During the 3D acquisition, all *z*-positions for each stained section were acquired. The system provides online analysis, right after acquisition, where it is decided whether to store the acquired image or not as well as offline analysis, which is performed on the stored images at a later time.

During the online analysis, the quality of the freshly acquired image is checked by computing the intensity histogram and applying the bilevel histogram analysis (17). There are two peaks in the intensity histogram: one corresponding to the background and the other corresponding to the objects. The two maxima are determined and the optimal threshold is computed as the local minimum between the two maxima. For the 2D analysis, the system first segments the nuclei using thresholding, where the authors compute the optimal threshold using the bilevel histogram analysis. Although a constant threshold is found sufficient for nuclei segmentation, the obtained binary image needs some improvement. To improve the binary image, they use morphological features, including nucleus size, presence of holes, nucleus roundness, and smoothness of boundaries. An output of the segmentation procedure is shown in Figure 3A. Then, the system detects the hybridized spots within each segmented nucleus based on a watershed-type technique called "gradual thresholding." The algorithm performs thresholding gradually from the highest threshold to the lowest one. The highest threshold is defined as the maximum stain intensity within the given nucleus. The lowest threshold is defined as the mean background intensity. For each step (threshold), thresholding-based segmentation is performed, and the result is compared with the segmentation result of the previous step. If a new object appears, it is considered to be a potential hybridization dot, unless it lies too close to an already existing dot. In this case, the two neighboring local maxima are considered as intensity fluctuations of one real hybridization dot.

Two alternative methods can be used for 3D image analysis. In the first, 2D analysis can be performed on the maximum image, which is defined as the image where each pixel's intensity is the maximum of the intensity of all the sections in that given lateral position. In the other method for analyzing the 3D images, all *z*-slices are analyzed using the 2D analysis and the information about the nuclei and the dots are stored. The corresponding nuclei and dots are found by comparing the output of each slice, by ensuring that the maximum intensity of each spot is always at the position of focus. As the authors mentioned, further improvements and optimizations of the proposed system can refine the evaluation of 3D images. However, the corresponding statistics are sufficient for a number of clinical and research tasks: routine diagnostics, follow-up of therapy, studies of chromatin structure, and many other different aspects of cell research.

Gue et al. (19) developed software, dubbed 3D FISH, to automate the spot segmentation and distance measurements in images from 3D FISH experiments. In their approach, all image processing is performed on voxels and a 3D color image is considered as a set of a potentially unlimited number of 3D gray-level images, each corresponding to a different color channel and probe.

The first step in image processing is the application of the median filter in order to remove the background noise. Then the 3D top-hat filter was employed to enhance the spots against the background. This second step of the algorithm is performed on all the different stacks, except for the DAPI channel. Spot detection is based on the consideration that pixels, whose intensity value is greater than the 99.95%, are determined as the central pixels of each spot. A local threshold was computed corresponding to the sum of the mean and the standard deviation of the intensity of the pixels belonging to three lines in the three directions (x, y, z) passing through the detected center of the spot. The seed is extended by a 3D connectivity to adjacent pixels whose value is greater than this local threshold to form an object. If the final object is touching one border of the image in the x-y plane, then it is removed. A different approach was employed for the DAPI channel. A median filter with a neighborhood of radius equal to 4 pixels is applied and then the ISODATA algorithm (8) is used for thresholding

the stack. At the end of both spot detection and DAPI processing, mathematical morphology procedures were applied to remove small objects, fill holes, and make shape look more compact. For validation, the authors performed a pilot 3D FISH study, producing 98% sensibility and 99% specificity for segmentation.

Kajtar et al. (20) applied a commercially available automated microscope slide scanning system (Metafer4, Metasystems, Germany) to detect the t(9;22)(q34;q11) in interphase nuclei of peripheral blood leukocytes. The automated analysis is separated in two steps: FISH signal detection and cell nucleus segmentation. The top-hat transform is used as the first step in signal spot detection. FISH signal recognition was accomplished using optimal values for the following parameters: the spot measurement area, minimum relative spot intensity, and minimum distance between two adjacent signals in the same color channel. These values were established by interactive training of the system. Nuclei segmentation was based on a fast contour following algorithm, where overlapping nuclei forming large clusters, small debris, and incomplete nuclei from the margins of the image field were removed. As in the spot detection procedure, optimal values for the maximum and minimum nuclear area, maximum concavity depth, and maximum aspect ratio were used to discriminate a single nucleus from the other objects. The data were analyzed manually by three independent investigators and then the results were compared with the automated ones. Although the false positive and false negative rates based on individual cells are not lower than that of manual analysis, interobserver variability is avoided using automated analysis, leading to increased statistical accuracy.

Lerner et al. (18,21–23) developed an approach for automatic FISH image analysis consisting of several ingredients, where the main advantage of their approach is its independence from auto-focusing mechanisms. In their method, the three color channels of the RGB (red-green-blue) image are analyzed separately. Emphasis is also placed on feature selection, using independent families of features, e.g., size, shape, intensity, color, and features projected on the data principal axes. Classification is based on hierarchical NNs (neural networks) partitioning of the feature space sequentially (21) as well as Bayesian methods (23).

Preprocessing of FISH images is performed in the RGB color space, while the HSI color space is used for multispectral image processing. Color segmentation using global thresholding is performed separately on each of the three different channels of RGB image. Following thresholding, the authors implement morphological operations in order to reduce the noise and to smooth the boundaries of the nuclei. After segmentation several features including eccentricity, area and spectral features, such as maximum and average hue, are extracted for each of the candidate signals. The authors try to classify the patterns (signals) into four classes: real red, artifact red, real green, and artifact green, following three classification strategies. In the first strategy, called the monolithic, patterns are classified into four classes using a single neural network. In the second, termed the independent strategy, patterns are classified into red and green classes using the "color network" and independently by a second network, the "real network," into real dots and artifacts. Classification of a pattern into one of the four classes is achieved by a common decision of both networks. In the third strategy, called the combined strategy, patterns are first classified into red and green classes using the color network. Then, based on the results of this network, they are classified by two other networks, the real-red network and the real-green network, into real and artifact red or green. In a later study (18), the same authors utilized the naive Bayesian classifier instead of neural network, to avoid dependency on a large number of parameters and neural network architecture settings, since the probability densities are the only parameters of the naive Bayesian classifier. Densities were evaluated by three methods: single Gaussian estimation which is parametric method, Gaussian mixture model assuming spherical covariance matrices, which is semiparametric method, and kernel density estimation, which is nonparametric method.

The evaluation of the three NN strategies and naive Bayesian classifier based on a database of 400 in-focus and out-focus images indicates high accuracy. The different maximum likelihood and maximum posterior solutions by the finite samples produce the slight difference in the performance of the Bayesian neural network and neural network. Although the naive Bayesian classifier is very simple, it is inferior compared with the other techniques. The inferiority of the naive Bayesian classifier is attributed to the conditional dependence of the features and to the additional inherent feature extraction stage of the NN classifiers.

Chawla et al. (24) developed an automated system for analyzing FISH signals from brain hippocampal and cortical sections. The authors employed several algorithms for automated 3D cell nuclei segmentation and FISH quantification. These algorithms have been organized and combined under a graphical user interface (GUI), referred to as "3D-catFISH." Confocal images are first obtained in multiple spectral channels and loaded onto the computer. The nuclear channel represents the counterstained nuclei and is first segmented using the 3D watershed algorithm, followed by a model-based region merging of the nuclei. Any of the cell types such as glia, which are to be excluded from the analysis are then removed by an agglomerative clustering algorithm. Intranuclear and cytoplasmic FISH signals are then detected and quantified using the other fluorescence channels. Finally, the measurements of nuclear segmentation, intranuclear and cytoplasmic FISH are integrated and associated spatially and a detailed tabular representation is generated.

During nuclei segmentation, the authors confronted the problem of the tight packing of cell layers which often results in the appearance of overlapping objects in the image stacks. The authors developed a combined image transform dubbed "gradient-weighted distance transform,"

which combines object separation hints from geometric and intensity cues in the images. An output of the proposed method is shown in Figure 3B. To correct for oversegmentation, it is necessary to detect and break the false watershed surfaces and thereby merge cell objects. This was accomplished via a model-based object merging method, where the object model was denoted by a vector, which includes various measurements such as volume, texture, convexity, circularity, and shape. Finally, unwanted glial/nonneuronal cells that normally appeared on the processing image were removed by a clustering algorithm using intensity, texture, and homogeneity features. The classification of FISH signals was performed by a computationally efficient algorithm, which identifies whether identifiable FISH products are present in the cell nucleus or in the cytoplasm. First the algorithm segments cell nuclei in the 3D image stack, then nonneuronal cells are removed. FISH signals are then quantified and cells are classified based on the presence of intranuclear or extranuclear/cytoplasmic FISH signals. It is reported that the validation results show 96.5% concordance with the human expert consensus.

O'Sullivan et al. (25) developed a method to assess the length of telomere in FISH images from cultured cells and human tissues of the gastrointestinal tract. The authors used the Optimas image analysis software (Media Cybernetics, Silver Spring, MD) to perform image analysis. The segmentation of the DNA image plane was based on the watershed algorithm and the identified nuclei were manually distinguished to either epithelial or stromal groups. Three methods were used to analyze the green (telomere) and in some cases the red (centromere) fluorescence intensity within each nucleus.

The first method, dubbed background corrected fluorescence, is based on the comparison of the difference between the intensity of the brightest and the dimmest pixels in each nucleus. The mean intensity of the dimmest 20% pixels, which is considered representative of nuclear background, is subtracted from the average of the brightest 5, 10, or 20% of red or green pixels in every nucleus. The resulting integer is the absolute telomere intensity. In the second method, which involves spot-finding, the telomere (green) or centromere (red) spots are identified using the watershed algorithm. The average intensity of all nonspot pixels within a specific nucleus is assumed as the background green or red fluorescence and using this value, the background corrected green or red spot pixel intensities are produced by subtracting that value from each pixel within a spot. In the third method, dubbed the background curve subtraction, authors create the histogram of pixel intensities for each nucleus, where the peak of the nuclear background fluorescence is considered as the mode of the green (or red) histogram. The mode is fine-tuned by fitting a second-order polynomial to the histogram in the region of the mode. Then a modification of the SFIT algorithm (26) is employed by reflecting the intensity histogram on the left side of the fitted mode and subtracted it from the histogram distribution on the right side of the mode. Finally, the authors calculate the integrated total integrated pixel intensity and average pixel intensity of the subtracted histogram.

The comparison of the three methods conclude that the spot-finding method is the more computationally demanding and has the greater variability, while background corrected fluorescence performs better on tissue sections. As the authors report, the measurements extracted with the proposed analysis methods are accurate, reproducible, and clinically applicable.

Based on the potential of further development of systems for the automated case-based reading of FISH images, Raimondo et al. (27), proposed a multistage algorithm for the automated classification of FISH images from breast carcinomas. The proposed algorithm can combine results from multiple images taken from a tissue slice to correctly classify the case. Despite the fact that the main content of FISH image red and green channels consists of spots, many FISH images frequently contain noisy areas consisting of large stains. For this reason, the authors started the spot detection with a top-hat filtering. Then a binary threshold was applied to the two output channels. Even the best threshold choice is not enough to isolate all true spots from false ones using only the red or green channel intensity. Thus, a spot template is computed for each channel. To measure the similarity between every candidate spot and the spot template, a cross correlation is used. Finally, for every detected spot, a channel intensity contrast measure is used to discard spots whose shape is very similar to the template one, but have a low channel intensity contrast with respect to their surrounding pixels, making them appear invisible to the human eye. The output of the spot detection algorithm is shown in Figure 4A.

Cell nuclei segmentation is performed on the FISH image blue channel. For many images, cell nuclei contain inhomogeneous blue channel intensities. To reduce the gray-level between dark regions and more illuminated ones, a nonlinearity correction step is performed. Then the algorithm by Otsu is employed to determine the threshold for initial nuclei segmentation. The binary image resulting from thresholding sometimes contains holes even at a single nucleus body region. This kind of holes has to be filled to enable correct nuclei segmentation. On the other hand, holes present in internuclei zones of overlapping nuclei should not be filled. To separate the two types of holes, the percentage P of the perimeter pixels of a circle centered on every hole centroid is calculated. After some experiments, the value for P varied in the range of 40–90% for holes of the first type and second type, respectively. The last step of the nuclei segmentation algorithm involves the marked watershed transform, which is employed to detect borders in overlapping nuclei clusters. In this step, the distance transform is first applied to the binary image obtained from the previous step. To reduce the number of spurious local maxima on the distance transform output, h-dome maxima is calculated. Figure 4B shows the result of the nuclei segmentation algorithm.

The authors used ROC curves to evaluate the performance of the proposed method, both for the spot detection

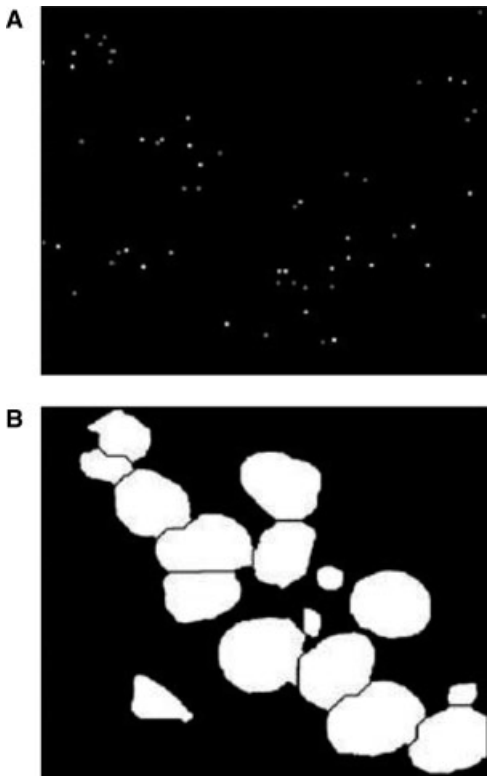

FIG. 4. Multistage algorithm for the automated classification of FISH images. (A) Output of the spot detection algorithm. (B) Final output of segmentation algorithm. Reprinted from Raimondo et al. (27), with permission from IEEE.

and nuclei segmentation on some FISH cases. Moreover, the overall algorithm performance for case-based classification on these cases showed the ability of the proposed system to distinguish between all positive and negative cases.

ANALYSIS OF IHC IMAGES

Aperion Technologies (28) have developed an open source software for IHC image analysis. The software includes three algorithms for four. The first algorithm, dubbed positive pixel count, calculates the area and the intensity of staining and assigns slides to four categories, negative, weak, medium, and strong. The algorithm only works on binary images, where the HSI color model is used to divide the color space into two classes, positive and negative. Thresholds are used to further divide the positive color class into three intensity ranges, which results in four total categories of staining. The intensity is calculated as the average of RGB pixel values and each pixel is assigned to one of the four categories. Thus, the total pixel count for each category is calculated.

The second algorithm, dubbed the nuclear algorithm, is used to count and measure nuclei in brown-positive and blue-negative IHC stained slides, where the most typical application for this algorithms is determination of the percentage of positive nuclei. The algorithm can also calculate the average size and staining intensity of each slide. Positive (brown) and negative (blue) regions are segmented based on color and morphological operators are used to further identify individual nuclei. The detection of object boundaries can be achieved using thresholding. For manual thresholding, a maximum and minimum intensity are selected to limit the range of valid intensities, while for automatic thresholding, amplitude and edge statistics are used to determine the range of intensity values that belong to the stained nuclei. Nuclei are discriminated by size and shape (roundness, compactness, and elongation), ignoring connective tissue and other objects.

The third algorithm, dubbed the membrane algorithm, is used to quantify the intensity and completeness (percentage) of membrane staining in IHC-stained slides, such as Her-2/neu. An important advantage of the algorithm is the ability to connect membranes that are not completely stained, in order to be able to score and classify the staining effect in every cell. The segmentation of cell membranes and nuclei is based on the watershed transform. Cells with size less than a threshold are not included in the final result. As a result, every cell contains one nucleus and is surrounded by a membrane. The average intensity and completeness of membrane staining is calculated for each cell and the cells are sorted based on the completeness of staining. The value of completeness that separates the top 10% of the cells was named as the completeness of staining.

The correct selection of thresholds is the most important part of the first and second algorithm. In the case of the first algorithm, the selection process is not fixed and remains undefined due to the nature of the problem. The lack of scoring-classification system is also a serious problem for the first algorithm. The membrane algorithm offers the ability to determine the completeness of membrane staining for each and every tumor cell. This is very important in order to interpret the result, as the scoring/classification system for the IHC effect evaluation highly depends on the percentage of membrane staining.

Weaver and Au (29) employed grayscale thresholding and size to segment nuclei from IHC slides stained from the proliferating cell nuclear antigen (PCNA) or bromodeoxyuridine (BrdUrd) labeling indices in human solid tumors using chromogen diaminobenzidine (DAB) and counterstained with hematoxylin. The nuclei were recognized as the pixels with low gray-level values (brown or blue), while the background and cytoplasm as the pixels with high gray level values. Thresholding for size was also applied, rejecting segmented nuclei outside a certain range. The color of staining, a basic element to score the immunohistochemical result, was utilized by transforming images from RGB format to the HSI (hue, saturation, intensity) model. Brown staining was distinguished from the blue counterstain using hue thresholds. To identify the extent of staining a minimum percent brown number (MPB) was defined. This value was the number of pixels labeled as brown in the entire recognized membrane boundary. The brown color is made up of magenta, yellow, and red hues while the blue color of only blue hues. As a result, hues representing brown staining are located

at the two ends of the hue spectrum and are separated by blue hues. Distinguishing the brown and blue colors requires two thresholds HT_1 and HT_2, where HT_1 separates the magenta from the blue hues and HT_2 separates the yellow and red hues from the blue hues. Pixels with values greater than HT_2 or less than HT_1 are considered brown, whereas pixels with values between HT_1 and HT_2 are considered blue. A typical value for HT_2 is 120, while MPB and HT_1 can be found by a trial and error procedure by first setting HT_1 to identify the brown stain and then setting MPB to correctly classify the cells as labeled or not. The trial and error procedure terminates when the identified cells agree with those identified by visual inspection. Automatic selection can be performed using the Otsu threshold-selection algorithm (29), which separates the histogram into two regions and defines one threshold, which divides only a portion of the hue histogram into two classes. The results of this study indicate that the accuracy of data results have improved using automatic thresholding in image analysis. However, the key point for successful selection of threshold is the high contrast between objects and background. Errors due to automatic thresholding alter the size and shape of objects but do not affect the number of determined objects.

Kim et al. (30) used an automatic video-color-image analysis system to study the immunostaining features of the androgen receptor (AR) in large samples of prostatic nuclei. Initially, the color images are converted to grayscale. Each grayscale image is binarized automatically by applying an adaptive thresholding algorithm. The segmentation results depend largely on the selected window size, where 80 pixels is a typical value, four times larger than a typical nucleus. A small window size will render the segmentation procedure faster but, in this case, the method will not recognize large nuclei. Nuclear masks are generated using the logical OR operation of three binary images, and their boundaries are smoothed by opening and closing with a 5×5 structuring element. Any remaining artifacts are eliminated by size and shape criteria. To measure the relative concentration of AR in immunopositive cells, the mean optical density (MOD) value is calculated by:

$$\mathrm{MOD} = -\frac{1}{N}\sum_{\mathrm{i}=1}^{N}\log\left(\frac{I_{\mathrm{i}}}{I_{\mathrm{O}}}\right) \qquad (2)$$

where N is the total number of pixels in a nuclear mask, I_{i} is the intensity level of pixel i, and I_{o} is the intensity level of the background measured in each field of view. The amount of MOD contributed by hematoxylin in the AR-stained section is corrected for by subtracting the average MOD of nuclei in the adjacent section. This corrected MOD is used as a measure of relative concentration of AR in immunopositive cells. An example of the nuclei segmentation procedure is presented in Figure 5. As reported by Kim et al., the detection of AR based on MOD in castrated mice is lower than in testosterone-stimulated mice, which indicates an initial success in qualitatively assessing the status of AR.

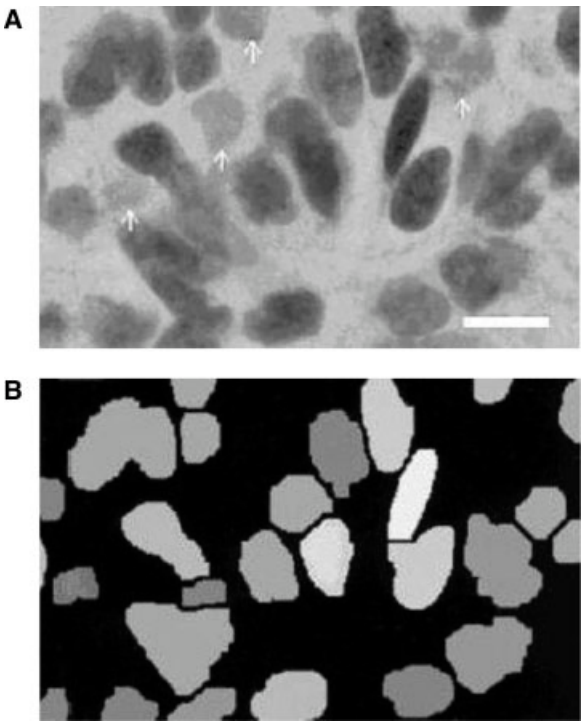

FIG. 5. Red, green, and blue parts of a color image of a sample field of view with AR immunostaining and hematoxylin counterstaining (A) were segmented using an adaptive thresholding algorithm and added to produce a set of binary masks for nuclear areas. Immunonegative nuclei (arrows) were removed using a linear discriminant analysis of their hue, saturation, and intensity values. MOD values were calculated and assigned to the nuclear mask as pixel intensity (B). Reprinted from Kim et al. (30), with permission from Wiley-Liss, John Wiley & Sons, Inc.

Another technique for IHC analysis was introduced by Smith et al. (31). The goal of this method is to quantify the expression level of a beta-gal antigen in neural tissue stained with immunofluorescent methods by determining the brightness-area-product (BAP). Using commercially developed software, a region of interest is selected from the test image and BAP is calculated, as the product of the number of pixels in a range with the difference between the mean and minimum brightness in the same region. An input and output threshold is defined for this region as the level of pixel brightness, above which background is unlikely to be found. To determine this threshold, the authors determine the range of brightness values occupied by no primary control images and define the upper limit of this range, above which any pixel of greater brightness intensity is likely to represent signal attributable to the epitope of interest, where the epitope is defined as the part of a macromolecule that is recognized by the immune system, specifically by antibodies, B cells, or cytotoxic T cells. Pixels of equal or greater brightness are then accepted. Threshold selection is based on histogram export of image pixel brightness.

By setting the minimum brightness equal to zero, the authors calculate the integrated optical density (IOD)

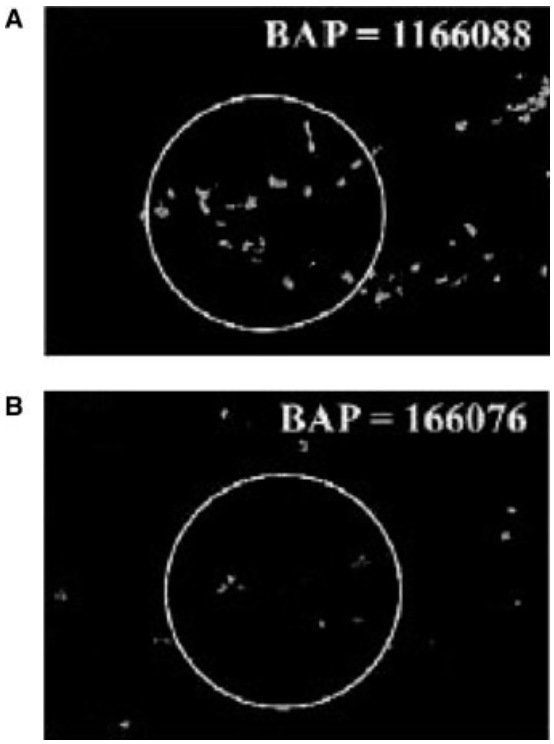

FIG. 6. Image taken from Smith et al. (31). (A,B) exemplify use of BAP in the setting of an alternative epitope distribution. In this case, 5′-bromo-2′-deoxyuridine (BrDU), a nuclear associated antigen which is more abundant in A than B by manual counting, is more abundant to a comparable extent as measured by BAP. Reprinted from Smith et al. (31), with permission from Elsevier.

which is used to cut the unlikely background. This describes the integrated brightness of the total area exceeding the minimum threshold for brightness of interest. BAP data was exported into a statistical software package for further analysis. The authors performed nonparametric ANOVA (distribution-free-Wallis test) followed by Dunn's test, for multiple comparisons, by using the Stata 7 intercooled™ software. The evaluation of this method suggests that it can be used to compare the staining in identical regions but cannot be used to count the amount of antigen present in an area. Moreover, the authors noticed versatility with regard to epitope distribution, as shown in Figure 6. Comparing the results with human observation, the BAP technique can be used as a simple intuitive measure of image brightness.

Kaczmarek et al. (32) employed very simple and objective quantitative techniques for IHC interpretation based on computer-assisted microscopy. The authors employ two approaches for segmentation of immunohistochemical reactions. The first approach is based on grayscale thresholding, where color images are first converted to grayscale and are enhanced with the median filter. The interval of grey shades, corresponding to the reaction, is defined and the area occupied by the reaction is extracted. Thresholding is applied and converts the foreground pixels into black color and the background pixels into white color. Focusing on the binary image, the area of positive reaction is calculated by counting the number of black pixels and the effect of the reaction by counting the percentage of black pixels in the image.

The second approach is based on spatial visualization of color reaction, using software designed and programmed by Nieruchalska et al. (33). At the beginning, the images are converted from the RGB to the HSB (hue, saturation, brightness) color space and are processed as 3D images in that space by introducing the intensity of color reaction as the third dimension. To measure the color reaction, spatial images are linearly converted to 256 colors. Pixels in red and yellow colors correspond to brown shades and are used to assess the area, volume, and intensity of color reaction. The spatial representation of the reaction is con-

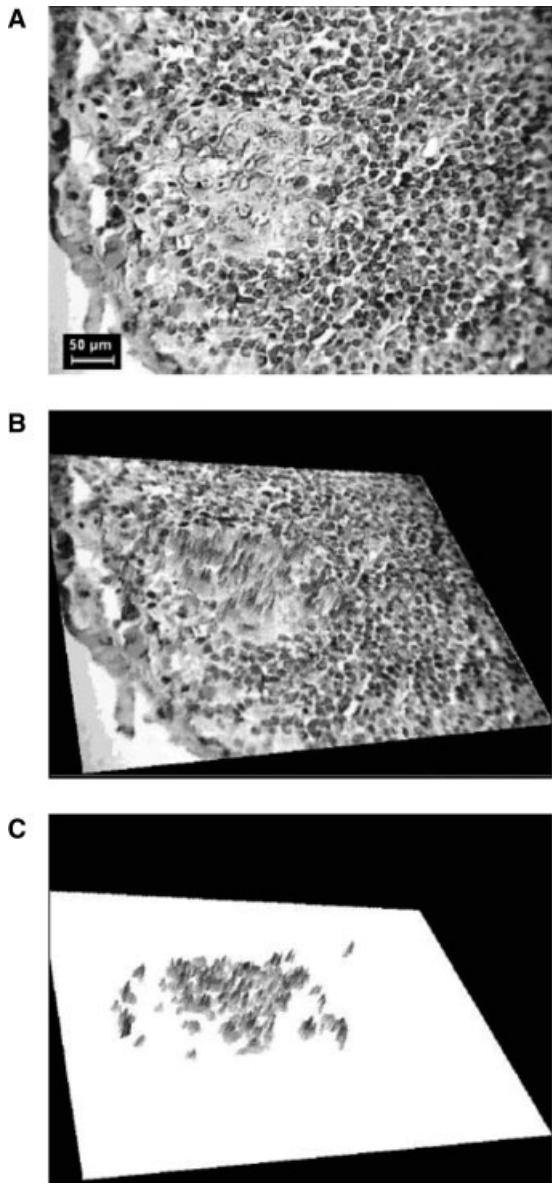

FIG. 7. (A) Micrographs of caspace-3. (B) Spatial visualization. (C) Segmented markers. Reprinted from Kaczmarek et al. (32), with permission from Akademia Medycznej w Białymstoku.

sidered as a set of connected prisms and pyramids to determine total volume the reaction. The reaction of the total area is derived from the orthogonal projection of the prisms and pyramids onto the plane and then the reaction intensity is derived from the volume/area ratio. Example of the proposed method applied on micrographs of caspace-3 is shown in Figure 7. The main advantage of this method is faster image analysis, up to 50 images per hour, including visual control of each image. The analysis can be more objective by using the same filters of colors, brightness, and saturation for a specific sequence of images.

Another approach is introduced by Matkowskyj et al. (34), in a study focused on determining the absolute amount of chromogen present by calculating the cumulative signal strength of IHC images. This was done by calculating the energy of images captured in Photoshop and processing the image file using Matlab. This study was focused on the gastrin-releasing peptide receptor, which is aberrantly expressed by human colon cancers. Human colon cancers variably express this receptor as a function of tumor differentiation, so this tissue type represents a good model for demonstrating the enhanced power of digital quantitative IHC. It is assumed that all information from the primary antibody-exposed slide contained in the experimental image file, excluding the information from the slide not exposed to primary antibody, but otherwise treated identically, is of importance. In this manner all the numerical data encoded within the file describing each pixel are used.

At first, the area of interest is selected for the control and the experimental image. The data contained within each RGB image file are stored as an $m \times n \times 3$ matrix, thus the "norm" function can be used to determine the singular value of the matrix, which represents the magnitude (mathematical energy) of the image file. The energy of an image is defined as:

$$E = \sqrt{\sum_{n_1=1}^{N_1}\sum_{n_2=1}^{N_2}\left(f_{\mathrm{RED}}(n_1,n_2)\right)^2} + \sqrt{\sum_{n_1=1}^{N_1}\sum_{n_2=1}^{N_2}\left(f_{\mathrm{GREEN}}(n_1,n_2)\right)^2} + \sqrt{\sum_{n_1=1}^{N_1}\sum_{n_2=1}^{N_2}\left(f_{\mathrm{BLUE}}(n_1,n_2)\right)^2} \quad (3)$$

where n_1, n_2 represents the pixel position in the $N_1 \times N_2$ image and f_{RED}, f_{GREEN}, and f_{BLUE} are the intensities of the red, green, and blue channels, respectively.

In general, algorithms based on pixel counting cannot determine the absolute amount of chromogen present, while algorithms based on color thresholding followed by an enumeration of the number of pixels present within a color/brightness range presuppose that the only information worth evaluating exists within a predefined spectral range. According to this algorithm, the cumulative energy of the control image is calculated and subtracted from that of the experimental image. To calculate the chromogen-specific energy, the energy of the region contained in the control image is subtracted from the homologous region of the experimental image:

$$E_{\mathrm{chromogen}} = E_{\mathrm{experimental}} - E_{\mathrm{control}} \quad (4)$$

Visual scoring of AR status of nuclei as positive or negative does not take into account the large variation in immunostaining intensity within a specimen or between specimens.

This method reduces the observer bias during the evaluation of specific cell regions. As mentioned, it can calculate the amount of chromogen precisely, by determining the energy resident in a number of tissue regions containing only cytoplasm. Therefore, it can be used to evaluate the nonhomogeneous tissue specimens. Another advantage is that the method is not restricted to evaluation of DAB-based IHC and it can quantify chromogen regardless of the type of secondary antibody detection system.

Nabi et al. (35) introduced an alternative method using image analysis to measure AR staining intensities by a pattern oriented approach: receptogram analysis. A receptogram is a composite of the univariate distribution of nuclear contents and their bivariate contour plots. In a receptogram, the IOD is plotted on the abscissa as a measure of the relative receptor content of individual nuclei and the percentage of nuclei expressing the different amounts of receptor were plotted on the ordinate. Based on contour slopes, the nuclear contents are classified into subtypes and each is correlated with response to treatment. Receptograms are classified into four categories, depending on the receptor immunostaining IOD distribution pattern. The first type includes receptograms with unimodal AR positive distributions, the second, with a bimodal distribution, the third, with multimodal AR positive distributions, and fourth, with a highly skewed distribution. The authors compare their results with the most common system used in the USA to grade the appearance of prostate cancer tissue (Gleason grade) and found no correlation. Thus, the method needs further studies before final conclusions can be drawn.

DISCUSSION

A high-level review of several methods of FISH and IHC image analysis from various tissue samples and research groups has been presented here. Analysis of FISH images is separated in two individual parts: spot detection and nuclei segmentation. Nuclei segmentation is performed in most of the algorithms; however, there exist examples where this step is not required. Regarding image processing algorithmic details, most methods employ an initial image quality enhancement step, mainly for shading correction and noise removal. This step is usually based on some type of morphological filter. Gray-level thresholding based on histogram analysis is also quite commonly encountered, both for nuclei segmentation and spot detection. The top-hat transform proved particularly useful for spot detection, while the distance and watershed transforms were prevalent for nuclei segmentation. Color information was also frequently utilized, either in the RGB

or HSI color spaces. There exist other approaches that were not extensively reviewed here since they do not focus on the image processing aspects of FISH image analysis. These include Narath et al. (36), who performed automatic telomere length measurement in interphase nuclei using the fluorescence-based automatic microscope (FLAME) and Wang et al. (37) who experimented with normalization of multicolor FISH (M-FISH) images to improve color karyotyping.

For IHC image analysis, most algorithms use color information from the HSI space, applying thresholding to identify staining, which is usually brown. Nuclei and cell segmentation is not as essential as in FISH; however, it is applied in some methods utilizing either the watershed transform or color information. Intensity and area of staining are also used to determine the extent of staining in most of the algorithms. This can be either in the form of image energy, or optical density, or the brightness area product.

The current status of most of these methods is at a point where large clinical studies are required to validate their effectiveness. This is especially critical, since most of the methods reviewed here report some deviations from the ground truth, as determined by the medical experts. The advantages of these trials are twofold, providing further hindsight for improving the algorithms as well as bringing in the medical experts as active participants of the development process.

As a conclusion, this review has highlighted the increasingly important role of automated image analysis methods in FISH and IHC, two very important diagnostic methods in medical practice. Manual intervention is still necessary to resolve particularly challenging or ambiguous cases, as well as to provide high-level supervision of the automatically-produced diagnostic results. It is envisioned that, with further algorithmic improvement of the automated methods, the role of manual intervention will decrease, thus increasing the diagnostic throughput and accuracy of FISH and IHC techniques. In addition, large-scale clinical trials will aid this procedure as well as lend more credibility to automated analysis methods.

ACKNOWLEDGMENTS

We thank Dr. Dourou at the Aristotle University of Thessaloniki for providing an example of an IHC image.

LITERATURE CITED

1. Kalliomeni OP, Kalliomeni A, Kurisu W, Thor A, Chen LC, Smith HS, Waldman FM, Pinkel D, Gray JW. ERBB2 amplification in breast cancer analyzed by fluorescence in situ hybridization. Proc Natl Acad Sci USA 1992;89:5321–5325.
2. Shapiro DN, Valentine MB, Rowe ST, Sinclair AE, Sublett JE, Roberts WM, Look AT. Detection of *N-myc* gene amplification by fluorescence in situ hybridization. Diagnostic utility for neuroblastoma. Am J Pathol 1993;142:1339–1346.
3. Ross JS, Fletcher JA, Bloom KJ, Linette GP, Stec J, Symmans WF, Pusztai L, Hortobagyi GN. Targeted therapy in breast cancer: The *Her-2/neu* gene and protein. Mol Cell Proteomics 2004;3:379–398.
4. Klijanienko J, Couturier J, Galut M, El-Naggar AK, Maciorowski Z, Padoy E, Mosseri V, Vieth P. Detection and quantification by fluorescence in situ hybridization (FISH) and image analysis of *Her-2/neu* gene amplification in breast cancer fine-needle samples. Cancer Cytopathol 1999;87:312–318.
5. Fernandez JL, Goyanes V, Fernandez CL, Buno I, Gosalvez J. Quantification of C-ERB-B2, gene amplification in breast cancer cells using fluorescence in situ hybridization and digital image analysis. Cancer Genet Cytogenet 1996;86:18–21.
6. Netten H, Young IT, van Vliet LJ, Tanke HJ, Vroljik H, Sloos WCR. FISH and Chips: Automation of fluorescent dot counting in interphase cell nuclei. Cytometry 1997;28:1–10.
7. Netten H, van Vliet LJ, Vroljik H, Sloos WCR, Tanke HJ, Young IT. Fluorescent dot counting in interphase cell nuclei. Bioimaging 1996;4:93–106.
8. Ridler TW, Calvard S. Picture thresholding using an iterative selection method. IEEE Trans Syst Man Cyber 1978;8:630–632.
9. Haralick RM, Sternberg SR, Zhuang X. Image analysis using mathematical morphology. IEEE Trans Pattern Anal Mach Intell 1987;9:532–550.
10. Young IT, Roos R. Acuity: Image analysis for the personal computer. In: Gelsema ES, Kanal LN, editors. Pattern Recognition and Artificial Intelligence. Amsterdam: Elsevier Science; 1988. pp 5–16.
11. Ortiz-De-Solorzano C, Santos A, Vallcorba I, Garcia-Sagredo JM, Del Pozo F. Automated FISH spot counting in interphase nuclei: Statistical validation and data correction. Cytometry 1998;31:93–99.
12. Castleman KR. Color compensation for digitized FISH images. Bioimaging 1993;1:159–165.
13. Beucher S, Meyer F. The morphological approach to segmentation: The watershed transformation. In: Dougherty E, editor. Mathematical Morphology in Image Processing. New York: Marcel Dekker; 1992. pp 433–481.
14. Serra J. Analysis and Mathematical Morphology. New York: Academic Press; 1982. pp 436–437.
15. Vincent L. Morphological grayscale reconstruction in image analysis. Applications and efficient algorithms. IEEE Trans Image Proc 1992; 2:172–201.
16. Kozubek M, Kozubek S, Lukasova E, Mareckova A, Bartova E, Skalnikova M, Jergova A. High-resolution cytometry of FISH dots in interphase nucleus nuclei. Cytometry 1999;36:279–293.
17. Pratt WK. Digital Image Processing. New York: Wiley; 1991. p 597.
18. Lerner B, Clocksin WF, Dhanjal S, Hulten MA, Bishop CM. Automatic signal classification in fluorescence in situ hybridization images. Cytometry 2001;43:87–93.
19. Gue M, Messaoudi C, Sun JS, Boudier T. Smart 3D-FISH: Automation of distance analysis in nuclei of interphase cells by image processing. Cytometry A 2005;67A:18–26.
20. Kajtar B, Mehes G, Lorch T, Deak L, Kneifne M, Alpar D, Pajor L. Automated fluorescent in situ hybridization (FISH) analysis of t(9;22)(q34;q11) in interphase nuclei. Cytometry A 2006;69A:506–514.
21. Lerner B, Clocksin WF, Dhanjal S, Hulten MA, Bishop CM. Feature representation and signal classification in fluorescence in-situ hybridization image analysis. IEEE Trans Syst Man Cyber 2001;31:655–665.
22. Clocksin WF, Lerner B. Automatic analysis of fluorescence in-situ hybridization images. Proc BMVC 2000:666–674.
23. Lerner B. Bayesian fluorescence in situ hybridization signal classification. Artif Intell Med 2004;30:301–316.
24. Chawla KM, Lin G, Olson K, Vazdarjanova A, Burke SN, McNaughton BL, Worley PF, Guzowski JF, Roysam B, Barnes CA. 3D-catFISH: A system for automated quantitative three-dimensional compartmental analysis of temporal gene transcription activity imaged by fluorescence in situ hybridization. J Neurosci Methods 2004;139:13–24.
25. O'Sullivan JN, Finley JC, Risques RA, Shen WT, Gollahon KA, Moskovitz AH, Gryaznov S, Harley CB, Rabinovitch PS. Telomere length assessment in tissue sections by quantitative FISH: Image analysis algorithms. Cytometry A 2004;58A:120–131.
26. Dean PN. A simplified method of DNA distribution analysis. Cell Tissue Kinet 1980;13:299–308.
27. Raimondo F, Gavrielides Ma, Karayannopoulou G, Lyroudia K, Pitas I, Kostopoulos I. Automated evaluation of Her-2/neu status in breast tissue from fluorescent in situ hybridization images. IEEE Trans Image Proc 2005;14:1288–1299.
28. Olson AH. Image Analysis Using the Aperio ScanScope. Technical manual. Aperio Technologies Inc., 2006.
29. Weaver JR, Au JL. Application of automatic thresholding in image analysis scoring of cells in human solid tumors labeled for proliferation markers. Cytometry 1997;29:128–135.
30. Kim D, Gregory CW, Smith GJ, Mohler JL. Immunohistochemical quantitation of androgen receptor expression using color video image analysis. Cytometry 1999;35:2–10.
31. Smith PD, McLean KJ, Murphy MA, Wilson Y, Murphy M, Turnley AM, Cook MJ. A brightness-area-product-based protocol for the quantitative assessment of antigen abundance in fluorescent immunohistochemistry. Brain Res Brain Res Protoc 2005;15:21–29.

32. Kaczmarek E, Gorna A, Majewski P. Techniques of image analysis for quantitative immunohistochemistry. Rocz Akad Med Bialymst 2004;49:155–158.
33. Nieruchalska E, Strzelczyk R, WoYniak A, Zurawski J, Kaczmarek E, Salwa-Zurawska W. A quantitative analysis of the expression of a-smooth muscle action in mesangioproliferative (GnMes) glomerulonephritis. Folia Morphol 2003;62:451–453.
34. Matkowskyj KA, Schonfeld D, Benya RV. Quantitative immunohistochemistry by measuring cumulative signal strength using commercially available software photoshop and matlab. J Histochem Cytochem 2000;48:303–312.
35. Nabi G, Seth A, Dinda AK, Gupta NP. Computer based receptogram approach: An objective way of assessing immunohistochemistry of androgen receptor staining and its correlation with hormonal response in metastatic carcinoma of prostate. J Clin Pathol 2004;57: 146–150.
36. Narath R, Lorch T, Greulich-Bode KM, Boukamp P, Ambros PF. Automatic telomere length measurements in interphase nuclei by IQ-FISH. Cytometry A 2005;68A:113–120.
37. Wang YP, Castleman KR. Normalization of multicolor fluorescence in situ hybridization (M-FISH) images for improving color karyotyping. Cytometry A 2005;64A:101–109.